On The Use of Confusion Matrixes

Frans Coenen
Department of Computer Science
University of Liverpool
Sunday 12 May 2012

A confusion matrix is a table describing the outcome from a binary supervised learning process (in the case of unsupervised learning it is called a matching matrix). Some authors refer to such a matrix as a truth table, but strictly speaking this is not correct. The matrix can be arranged in a number of ways, one example is as shown in Table 1.

		Bench Mark		
		X	Not X	
Classifier	X (Positive)	True Positives (TP), records of class X that were correctly classified as X	False Positives (FP), records of class Not X that were incorrectly classified as X	Total Positive classifications
	Not X (Negative)	False Negatives (FN), records of class X that were incorrectly classified as class Not X	True Negatives (TN), records of class Not X that were correctly classified as class Not X	Total Negative classifications
		Total number of records belonging to class X	Total number of records belonging to class Not X	Total number of records

Table 1: *Example confusion matrix*

Given a confusion matrix of the form given in Table 1, we can generate various statistics. With respect to the columns we can calculate the *sensitivity* and *specificity*:

$$sensitivity = \frac{TP}{TP + FN}$$

$$specificity = \frac{TN}{TN + FP}$$

Sensitivity is also referred to as *recall* or the *true positive rate*. It is a measure of how well a classifier can be used identify classify instances as belonging to a certain class. Specificity corresponds to the *true negative rate*. Specificity is the same thing as sensitivity but for the negative class label. A good classifier is one that maximizes both sensitivity and specificity (we want our classifiers to be both sensitive and specific).

With respect to the rows we can calculate the *precision* with respect to a certain class:

$$precision_X = \frac{TP}{TP + FP}$$

$$precision_{NotX} = \frac{TN}{FN + TN}$$

Precision is a measure of how well a classifier performs with respect to a specific class. Again the higher the precision the better.

Finally we can consider the entire confusion matrix and calculate the *accuracy*:

$$accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

The F_1 measure is the "harmonic mean" of the precision with respect to X and sensitivity values (it is called the F_1 measure because precision and sensitivity are equally weighted, there are other weighted variations). The F_1 value is calculated as follows:

$$F_1 = 2 \times \frac{precision \times sensirivity}{precision + sensitivity}$$

Examples

We will now consider a number of examples. The results have been collected together in Table 2. Example 1 gives the confusion matrix for a "perfectly accurate" classifier (everything classified correctly). Example 2 gives the confusion matrix for a "perfectly inaccurate" classifier (everything classified wrongly). In the first case sensitivity, specificity, precision and accuracy are all one; in the second case they are all zero.

		Bench Mark		
		X	Not X	
Classifier	X (Positive)	50	0	50
	Not X (Negative)	0	50	50
		50	50	100

Example 1: *Confusion matrix for a perfectly accurate classifier*

		Bench Mark		
		X	Not X	
Classifier	X (Positive)	0	50	50
	Not X (Negative)	50	0	50
		50	50	100

Example 2: *Confusion matrix for a perfectly inaccurate classifier*

Example 3 gives the case where everything is classified as X, and Example 4 the case where everything is classified as Y (thus a single default rule situation). The accuracy in both cases is 0.5 (50%). However, the sensitivity in the first case is 1.0 (all X cases correctly classified), but the specificity is 0.0. In the second case the reverse is true.

		Bench Mark		
		X	Not X	
Classifier	X (Positive)	50	50	100
	Not X (Negative)	0	0	0
		50	50	100

Example 3: *Confusion matrix all records classified as X*

		Bench Mark		
		X	Not X	
Classifier	X (Positive)	0	0	0
	Not X (Negative)	50	50	100
		50	50	100

Example 4: *Confusion matrix all records classified as Not X*

Supposing we now have a situation where half the records (examples) are classified correctly (Example 5). This is the situation where we are simply guessing. In this case sensitivity, specificity, precision and accuracy will all be 0.5. This given an equally distributed data set we want the sensitivity, specificity, precision and accuracy to be greater than 0.5.

		Bench Mark		
		X	Not X	
Classifier	X (Positive)	25	25	50
	Not X (Negative)	25	25	50
		50	50	100

Example 5: *Confusion matrix where records equally distributed (guessing)*

Just to finish off with, suppose we have a classifier that is good at identifying records belonging to class X, but not so good at identifying records that do not belong to class X (Example 6). We have a highly sensitive classifier and one that is very precise with respect to Class Not X (no records belonging to class X wrongly classified as belonging to class Not X), however the specificity is poor.

		Bench Mark		
		X	Not X	
Classifier	X (Positive)	50	25	75
	Not X (Negative)	0	25	25
		50	50	100

Example 6: *Confusion matrix for a classifier that performs well with respect to class X and not so well (half right) with respect to class Not X*

	Example 1	Example 2	Example 3	Example 4	Example 5	Example 6
Sensitivity	1.000	0.000	1.000	0.000	0.500	1.000
Specificity	1.000	0.000	0.000	1.000	0.500	0.500
Precision X	1.000	0.000	0.500	0.000	0.500	0.667
Precision Not X	1.000	0.000	0.000	0.500	0.500	1.000
Accuracy	1.000	0.000	0.500	0.500	0.500	1.000
F_1 Measure	1.000	0.000	0.667	0.000	0.500	0.800

Table 2: *Sensitivity, Specificity, Precision and Accuracy Values for Examples 1 to 6.*